Erdős's 1974 distance-product volume problem resolved: the ball as unique maximizer, sharp inradius stability, and other miscellanea

Dave Goel

First version: November 19, 2026. This version: November 19, 2026.

hexagon:2611.00701v1

Abstract

We consider the set of all points of $\mathbb{R}^d$ whose distances to n given points have product at most one. Erdős (1974) asks how large its volume can be. We show that in every dimension $d \geq 3$, and for every integer $n \geq 1$, this volume is at most that of the unit ball. Equality holds only and exactly when all n given points coincide. This completely answers Erdős's question; in the plane, the same statement is Pólya's theorem of 1928. In fact we settle a more general problem, with weighted points. Our theorem also shows that the two-point counterexample in dimension three, reported by Erdős in 1976 and repeated in 2018 and 2026, cannot exist. We also prove a sharp stability estimate: for each fixed $d \geq 3$, if such a set falls short of the maximal volume by a small fraction D, it still contains a ball of radius at least $1 - (1/(d-2) + o(1))D$ as D tends to zero, uniformly in the number, weights and positions of the points. The same holds for every compactly supported probability measure in place of the points. Two equally weighted points approaching each other show that the coefficient $1/(d-2)$ cannot be lowered. Finally, the product of the distances is at most one exactly when their geometric mean is at most one, and we ask what happens when the geometric mean is replaced by the weighted power mean of order p. We show that the same volume bound holds for all placements exactly when $p \geq 2 - d$, and that in this range equality again requires coincident points. For $p \geq 2 - d$ we deduce the bound from the weak estimate of Tang and Wang; for $p < 2 - d$ the polarization theorem of Erdélyi and Saff gives placements that violate it. We prove the equality statement at the endpoint $p = 2 - d$ directly, and for $p < 2 - d$ we compute the leading term of the excess volume of the region of two nearby points.

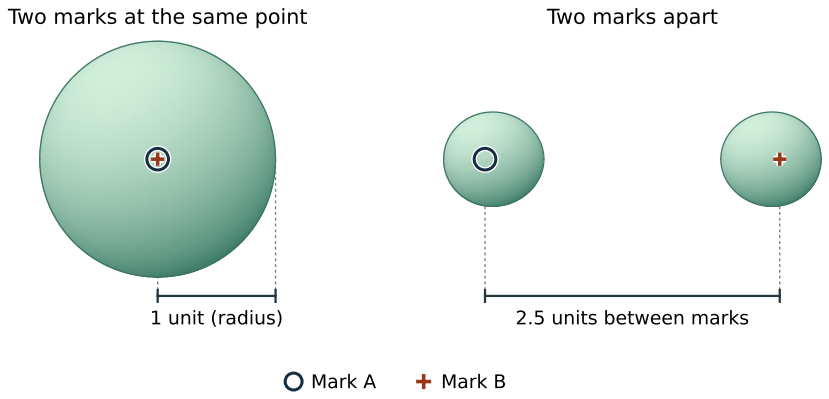
This is the expanded version of the paper "Erdős's 1974 distance-product volume problem resolved: the ball is always the unique maximizer in $d \geq 2$" [4].

The volume problem

Fix a dimension $d \geq 2$ and a number $n \geq 1$. Place n points $a_1, \ldots, a_n$, called marks, anywhere in $\mathbb{R}^d$; marks may coincide. Let E consist of the points x for which $|x - a_1| \cdots |x - a_n| \leq 1$. *How large can the volume of E be, and which placements of the n marks attain the maximum?* Erdős's question dates from 1974; it is Problem 4.23 in the 2018 list of Hayman and Lingham [5, Problem 4.23, Tables 1–2].

If all n marks lie at one point a, the product is $|x - a|^n$, and E is the closed unit ball about a. We write ω_d for its volume; thus $\omega_3 = 4\pi/3$. We prove that, in every dimension $d \geq 3$ and for every n, the largest volume of E is ω_d, and that every placement in which the marks do not all coincide gives strictly less volume (Theorem 1). In fact we settle a more general problem: Theorem 1 allows the marks to carry different weights. Erdős's problem is the case of equal weights.

Best result before this paper	**New result**
What is the largest volume of E, for n marks in dimension $d \geq 2$, and when is it attained?	
Known only for $d = 2$ (Pólya [12]). Listed as open for $d \geq 3$ [5, Problem 4.23], [6, abstract].	Fully resolved for all $d \geq 2$ and $n \geq 1$: the maximum is ω_d, attained exactly when all marks coincide.
Can two distinct marks in dimension three give more volume than the unit ball?	
Reported possible, incorrectly. Erdős wrote in 1976 that Piranian had a counterexample [3]. Hayman and Lingham repeated this in 2018 [5], and Ferudun in 2026 [6].	No. Every pair of distinct marks gives volume strictly less than $4\pi/3$. The reported counterexample does not exist.



	The two marks coincide	The two marks are distinct
Earlier status	Ball volume $4\pi/3$ known; is it maximal?	A larger volume was reported, incorrectly, for a distinct pair
Theorem 1	Yes: maximum volume $4\pi/3$	No: every distinct pair gives less

Figure 1: Two marks in dimension three, drawn to the same scale. Left: the marks coincide, and the region is the unit ball. Right: the marks are 2.5 units apart, and the region splits into two pieces whose combined volume is less than that of the unit ball. The separation 2.5 is our choice; Erdős's report gave no separation.

To see how the dimension enters the problem, take two marks at distance $2t$ from each other, with $0 < t \leq 1$, and let m be their midpoint. A point on the line through the marks, at distance s from m, has distances $|s - t|$ and $s + t$ to the marks, with product $|s^2 - t^2|$. This product is at most one exactly when $t^2 - 1 \leq s^2 \leq t^2 + 1$. Since $t^2 \leq 1$, the left inequality always holds, so the product is at most one exactly when $s \leq \sqrt{1 + t^2}$. So along that line E reaches to distance $\sqrt{1 + t^2}$ from m, which is more than one. A point at distance b from m, on a line through m perpendicular to the first, has distance $\sqrt{b^2 + t^2}$ to each mark, with product $b^2 + t^2$. So through the midpoint m the region E reaches sideways only to distance $\sqrt{1 - t^2}$, which is less than one. Separating the marks therefore makes the region longer along their line and thinner across it at the middle, in each of the $d - 1$ perpendicular directions. Whether the volume grows or shrinks as a result depends on the dimension.

- *The real line, $d = 1$.* For two marks, the greatest length is $2\sqrt{2}$, attained when the marks are two units apart. Indeed, for $0 \leq t \leq 1$ the region is a single interval of length $2\sqrt{1 + t^2}$. Its length increases from two when the marks coincide to $2\sqrt{2}$ when the marks are at -1 and 1.

For $t > 1$ the region splits into two intervals. Their total length is

$$2(\sqrt{t^2+1} - \sqrt{t^2-1}) = \frac{4}{\sqrt{t^2+1} + \sqrt{t^2-1}} < 2\sqrt{2}.$$

So separating the marks by more than two units makes the total length of the region smaller.

- *The plane, $d = 2$.* The unit disc is the best region. Pólya proved in 1928 that, for every n and every placement of the marks, the area of E is at most π, with equality only when all the marks coincide [12, Satz II, p. 280].
- *Every dimension $d \geq 3$.* The unit ball is the best region: by Theorem 1, the volume of E is at most ω_d, with equality only when all the marks coincide.

In three dimensions, Erdős at first expected the unit ball to have the largest volume [3, §4]. A single mark always gives a unit ball, so two marks are the first case in which the placement matters.

In 1976, Erdős reported that Piranian had found a two-point counterexample to the claim that the unit ball maximizes volume in three dimensions [3, §4, pp. 350–351]. Erdős gave neither the configuration nor a volume calculation. That report is incorrect: the counterexample cannot possibly exist. In the proof of Theorem 1, we show that every pair of distinct marks gives volume strictly less than $4\pi/3$, the volume attained when the two marks coincide. We know of no publication before the present paper that explicitly corrected Erdős's report. Hayman and Lingham repeated the report in 2018 [5, Problem 4.23, p. 79]. Ferudun repeated it in 2026, in a summary of Hayman and Lingham's Problem 4.23 [6, §1.1, p. 1].

Problem 4.23 asks a second question as well: whether E, when it is connected, lies in a ball of radius two. Ferudun answered that covering question, in every dimension [6, Theorem 1.1]. The present paper concerns the volume question only.

We state Theorem 1 in the more general form, with weighted marks. Give each mark a_j a weight $w_j > 0$, with the weights summing to one. The product of the distances is replaced by their weighted geometric mean G, which is compared with an arbitrary threshold $R > 0$. Erdős's region is the case of equal weights $w_j = 1/n$ and $R = 1$, because the product of n distances is at most one exactly when their geometric mean is at most one. Thus $E = E_1$ in the notation of the theorem.

Theorem 1. *Let $d \geq 3$, let $a_1, \ldots, a_n \in \mathbb{R}^d$ with $n \geq 1$, and let $w_1, \ldots, w_n > 0$ have sum one. For $R > 0$ put*

$$G(x) = \prod_{j=1}^{n} |x - a_j|^{w_j}, \qquad E_R = \{x : G(x) \leq R\}.$$

Write ω_d for the volume of the unit ball in $\mathbb{R}^d$. Then $\mathrm{vol}_d(E_R) \leq \omega_d R^d$, with equality exactly when the a_j coincide. In particular, for $d = 3$, $w_j = 1/n$ and $R = 1$, the maximum in Erdős's problem is $4\pi/3$.

For equal weights and $R = 1$, Theorem 1 answers Erdős's volume question in every dimension $d \geq 3$: the largest volume is ω_d, and it is attained exactly when all marks lie at one point.

We first treat two marks in dimension three, the case of the reported counterexample, by applying Pólya's planar theorem to each cross-section. We then give two proofs of Theorem 1: a direct proof by flux across level surfaces, which follows Talenti's rearrangement method [7], and a second proof that combines electrostatics and isoperimetric geometry. The later sections treat the largest inscribed ball, first in dimension three and then in every dimension $d \geq 3$. The final section replaces the geometric mean of the distances by a power mean.

Two marks in dimension three: a short proof

Let the two marks be distinct, and choose coordinates in which both lie in the horizontal plane of height zero. Take a point at height h, and let u and v be the distances from its vertical projection onto the horizontal plane to the two marks. The distances from the point itself to the marks are $\sqrt{u^2+h^2}$ and $\sqrt{v^2+h^2}$, and the square of their product is

$$(u^2+h^2)(v^2+h^2) = (uv+h^2)^2 + h^2(u-v)^2.$$

For a point of the region the left side is at most one. The right side is at least $(uv+h^2)^2$, so $uv+h^2 \le 1$, that is, $uv \le 1-h^2$. Hence the region has no point at a height h with $|h|>1$, and for $|h|<1$ its horizontal slice at height h lies in the planar region $uv \le 1-h^2$. Identify the horizontal plane with the complex plane, and let F be the monic quadratic polynomial whose zeros are the two marks. If z is the vertical projection of the point, read as a complex number, then $uv = |F(z)|$. Pólya's theorem [12, Satz II, p. 280] states that, for a monic polynomial P of degree n and every $R>0$, the region $|P(z)| \le R$ has area at most $\pi R^{2/n}$, with equality only when $P(z) = (z-c)^n$ for some complex number c. (In this statement n is the degree and R is a bound on $|P|$; R is not the threshold of Theorem 1.) For $P=F$, with $n=2$ and $R = 1-h^2$, the planar region $uv \le 1-h^2$ has area at most $\pi(1-h^2)$, and strictly less because the two marks are distinct. Pólya's bound holds for the whole planar region, including both pieces if the region is disconnected. It therefore holds for the whole slice. The slice of the unit ball at the same height has area exactly $\pi(1-h^2)$. Integrating the slice areas over $-1<h<1$ shows that the region has volume strictly below $4\pi/3$. Hence no two distinct marks give as much volume as the unit ball.

Erdős, Herzog and Piranian restate Pólya's theorem for the open region $|P(z)|<1$ [11, p. 134]. For the two marks used here, the same bound follows from that form: apply it to the monic quadratic $F(\sqrt{R}\,z)/R$, scale back, which multiplies the area by R, and note that the boundary $|F(z)|=R$ has area zero. Indeed, write $z = x_1 + ix_2$ with x_1 and x_2 real. For fixed x_1, $|F(x_1+ix_2)|^2$ is a monic quartic polynomial in x_2, so it equals R^2 at no more than four points of the vertical line through x_1. Hence the boundary has area zero by Fubini's theorem.

Why the separated marks in Figure 1 give two pieces

Write a point as $(s,y) \in \mathbb{R}\times\mathbb{R}^2$ and put the marks at $(\pm t, 0)$ with $t=5/4$. The square of the product of the two distances is

$$(s^2-t^2)^2 + 2(s^2+t^2)|y|^2 + |y|^4.$$

For fixed s this expression increases with $|y|$, so it decreases along the whole segment from (s,y) to $(s,0)$ whenever $y \ne 0$. On the axis the region occupies precisely $3/4 \le |s| \le \sqrt{41}/4$. Each point of the region can therefore be joined inside the region to the mark on the same side of the plane $s=0$, first by the transverse segment and then along the axis. No point of the plane $s=0$ belongs to the region, because there the expression equals $(t^2+|y|^2)^2 \ge t^4 > 1$. This plane separates the two halves.

First proof: flux across level surfaces

The idea is to enclose E_R in a larger set, a superlevel set of the Newtonian potential of the marks, and to show that the volume of this larger set is at most $\omega_d R^d$. When two marks lie at different points, the larger set contains an open ball that E_R misses, and this gives the strict inequality. At each mark a_j place a point charge of size w_j; from now on we also call the marks *charges*.

Write $U(x) = \sum_j w_j |x - a_j|^{2-d}$, and set $U = +\infty$ at each charge. The arithmetic–geometric mean inequality gives $U(x) \geq G(x)^{2-d}$ away from the charges. At a point of E_R that is not a charge we have $G \leq R$, and the exponent $2 - d$ is negative, so $U \geq G^{2-d} \geq R^{2-d}$. At a charge U is infinite. Hence $E_R \subset \{U \geq R^{2-d}\}$.

For $\lambda > 0$ let $W(\lambda) = \mathrm{vol}_d\{U \geq \lambda\}$ and define $r(\lambda) > 0$ by $W(\lambda) = \omega_d r(\lambda)^d$. Thus $r(\lambda)$ is the radius of the ball that has the same volume as the superlevel set $\{U \geq \lambda\}$. Each superlevel set is bounded and contains a neighborhood of each charge. At a regular value λ, let $A(\lambda)$ denote the area of its boundary, the level surface $\{U = \lambda\}$. Because U is harmonic away from the charges,

$$\int_{U=\lambda} |\nabla U| \, dS = d(d-2)\omega_d.$$

Indeed, merge charges that sit at the same point, and remove from the superlevel set a small ball about each of the resulting charges. On the level surface the outward normal derivative of U is $-|\nabla U|$. Through each small sphere, in the direction of the normal that points toward its center, the term of U that is singular at the center has flux $d(d-2)\omega_d$ times the size of the charge at that center, and every other term has flux zero. The sizes of the charges sum to one, so the divergence theorem gives the identity.

For $0 < \kappa \leq \lambda < \infty$, call the set where $\kappa \leq U \leq \lambda$ a *slab*. On every slab, U is smooth and its critical set has volume zero. To see this, note that U is analytic and nonconstant on the connected complement of the charges, so at least one of its partial derivatives is not identically zero there. That derivative is again analytic, so its zero set has volume zero; and this zero set contains the critical set. The same argument applied to $U - \lambda$ shows that each level set $\{U = \lambda\}$ has volume zero. The coarea formula on the slab, applied away from the critical set, shows that W is locally absolutely continuous and that $-W'(\lambda) = \int_{U=\lambda} |\nabla U|^{-1} \, dS$ for almost every λ. By Sard's theorem, almost every λ is a regular value. For almost every λ we therefore have

$$(d\omega_d r(\lambda)^{d-1})^2 \leq A(\lambda)^2 \leq d(d-2)\omega_d \, (-W'(\lambda)).$$

The first inequality is the isoperimetric inequality for the set $\{U \geq \lambda\}$, whose volume is $\omega_d r(\lambda)^d$. The second is the Cauchy–Schwarz inequality $A(\lambda)^2 \leq \int_{U=\lambda} |\nabla U| \, dS \cdot \int_{U=\lambda} |\nabla U|^{-1} \, dS$, combined with the flux identity and the formula for $-W'(\lambda)$. Substituting $W = \omega_d r^d$, so that $-W' = -d\omega_d r^{d-1} r'$, and cancelling gives $-(d-2) r^{1-d} r' \geq 1$, that is, $(r^{2-d})' \geq 1$ almost everywhere. As λ decreases to zero the superlevel sets exhaust $\mathbb{R}^d$, so $r(\lambda)^{2-d} \to 0$. Since W is locally absolutely continuous, so is r^{2-d}, and integrating from zero to λ yields $r(\lambda)^{2-d} \geq \lambda$. At $\lambda = R^{2-d}$ it follows that $r(\lambda) \leq R$. Since E_R lies in $\{U \geq R^{2-d}\}$, this proves the volume bound. The same bound on the superlevel set follows from Tang and Wang's sharp weak estimate for densities [8, Theorem 1.1, $s = 2$] by smoothing the point charges and passing to the limit; see the deduction after Theorem 4.

It remains to treat equality. Suppose first that two marks lie at different points a and b. Consider the ray that starts at a and points directly away from b. At every point of this ray the distances to a and to b are different. The compact set E_R contains a. Let y be the last point of the ray that belongs to E_R. By continuity, $G(y) = R$. Because the distances from y to a and to b differ, the arithmetic–geometric mean inequality is strict at y, so $U(y) > R^{2-d}$. A point on the ray just beyond y therefore has both $G > R$ and $U > R^{2-d}$, and by continuity both inequalities hold on a small open ball about it. This ball lies in the superlevel set but outside E_R. Thus E_R has strictly smaller volume than the superlevel set $\{U \geq R^{2-d}\}$, and hence volume strictly less than $\omega_d R^d$. If instead all the marks coincide at a, then $G(x) = |x - a|$ and E_R is the ball of radius R about a, of volume exactly $\omega_d R^d$. □

Second proof: combining electrostatics and isoperimetric geometry

We treat E_R and a ball of radius R as conductors. Using Clark and Laugesen's logarithmic moment inequality, we show that, when both conductors are held at unit potential, E_R carries at most as much charge as the ball. Newtonian capacity measures the charge that a conductor carries at unit potential. By the isocapacitary inequality, a region with at most the ball's capacity also has at most the ball's volume. The compact set E_R contains a ball and hence has positive Newtonian capacity. Let c be the radius of a ball of the same capacity and let ν be the Newtonian equilibrium probability measure of E_R, the distribution of one unit of charge on E_R that minimizes its Newtonian energy. For a ball, this equilibrium measure is the uniform measure on its boundary sphere. Applying Theorem 1.1(ii) of Clark and Laugesen [1], with the origin moved to a_j, gives $\int \log|x-a_j|\,d\nu(x) \geq \log c$. Multiplying by the weights w_j and summing over j gives

$$\log c \leq \int_{E_R} \log G(x)\,d\nu(x) \leq \log R.$$

The second inequality holds because $G \leq R$ on E_R. The classical isocapacitary inequality (see [2, §1]) gives $\mathrm{vol}_d(E_R) \leq \omega_d c^d$. That inequality is stated for bounded open sets. For a compact set, apply it to bounded open neighborhoods of the set and take the infimum of their capacities. Since Newtonian capacity is outer regular, this infimum is the capacity of the compact set. We use the resulting inequality for compact sets here and in the inradius argument below.

Since $c \leq R$, the isocapacitary inequality gives the volume bound $\mathrm{vol}_d(E_R) \leq \omega_d R^d$. For the equality case, suppose $\mathrm{vol}_d(E_R) = \omega_d R^d$. Then $c = R$, so both inequalities in the last display are equalities. Each inequality $\int \log|x-a_j|\,d\nu(x) \geq \log c$ enters the sum with a positive weight, so each of them is an equality as well. By Clark and Laugesen [1, Theorem 1.1(iv)], E_R contains the entire sphere of radius R about each a_j. Choose a unit vector u, and among the marks choose one, a, for which $u \cdot a$ is largest. The point $a + Ru$ is at distance R from a. For every mark b that lies at a different point we have $u \cdot (a-b) \geq 0$, and therefore

$$|a + Ru - b|^2 = R^2 + 2R\,u \cdot (a-b) + |a-b|^2 > R^2.$$

So if two marks lay at different points, the geometric mean $G(a+Ru)$ would exceed R. This contradicts the fact that E_R contains the sphere of radius R about a. □

The largest inscribed ball near maximum volume

Theorem 1 gives the exact maximum volume in every dimension $d \geq 3$. We now ask how large a ball must still fit inside a region whose volume is just below that maximum. In this section we prove the following, in three dimensions. If the volume of the region falls short of the maximum $\omega_3 R^3$ by a small fraction, then the region still contains a ball whose radius falls short of R by at most the same fraction, to first order. The coefficient one is sharp. Theorem 3 below extends the estimate to every dimension $d \geq 3$, where the sharp coefficient is $1/(d-2)$.

The finitely many weighted marks may be replaced by any compactly supported Borel probability measure μ on $\mathbb{R}^3$. We call μ a *distribution of marks*. The finite configuration in Theorem 1 corresponds to $\mu = \sum_j w_j \delta_{a_j}$, where δ_a is the unit mass at a, and G_μ below is then the weighted geometric mean G of that theorem. Set

$$G_\mu(x) = \exp\int \log|x-a|\,d\mu(a), \qquad E_R = \{x : G_\mu(x) \leq R\}, \qquad \exp(-\infty) = 0.$$

The integral $\int \log|x-a|\,d\mu(a)$ is the decreasing limit, as $M \to \infty$, of the continuous functions $\int \max(\log|x-a|, -M)\,d\mu(a)$. Thus G_μ is a measurable function and E_R is a measurable set. The set E_R is bounded because μ has compact support. Write ρ for the *inradius* of E_R, the supremum of the radii of balls contained in E_R, taking $\rho = 0$ if none fits. Write $|A|$ for the volume of a set A, and define the fractional volume deficit by $D = 1 - |E_R|/(\omega_3 R^3)$, where $\omega_3 = 4\pi/3$. Thus D is the fraction of the maximal volume that is missing, and $1 - \rho/R$ is the fraction of the maximal radius that is missing. At threshold one, the argument with compact subsets given below shows that $|E_1| \le \omega_3$. Rescaling does not change D. Hence $D \ge 0$.

Theorem 2. *For each $R > 0$,*

$$\lim_{h \downarrow 0} \sup_{\mu:\, 0 < D \le h} \frac{1 - \rho/R}{D} = 1.$$

The supremum ranges over all compactly supported probability measures in $\mathbb{R}^3$, with no common bound on the radii of their supports. The same limit holds if the supremum is taken only over finite sets of equally weighted marks.

Here h is the largest volume deficit allowed. The notation $h \downarrow 0$ means that we lower this positive bound toward zero. As the missing fraction D of the maximal volume shrinks to zero, the missing fraction of the maximal radius is at most about D, and the coefficient one in front of D cannot be lowered. In particular, given $\varepsilon > 0$, every such region with sufficiently small positive deficit satisfies $1 - \rho/R \le (1+\varepsilon)D$. Theorem 3 below gives the corresponding statement in every fixed dimension $d \ge 3$, with the coefficient $1/(d-2)$ in place of one.

Under the dilation $x \mapsto x/R$ of space and of the marks, every distance is divided by R. Thus G_μ and ρ are divided by R, and D stays fixed. We prove Theorem 2 at $R = 1$.

Two marks approaching one another

We begin with the example that shows why the coefficient one cannot be lowered. Move two equally weighted marks toward each other. Their region approaches the unit ball. We first find the largest ball that fits inside the region, and then compare the fractional loss of radius with the fractional loss of volume. Figure 2 shows one position along the way, with the marks 1.2 units apart.

Put mass $1/2$ at each of $\pm t e_1$, where $e_1 = (1,0,0)$ and $0 < t < 1/\sqrt{2}$. The squares of the two distances from a point x to the marks have arithmetic mean $|x|^2 + t^2$. The product of the two distances is the geometric mean of these squares, so it is at most $|x|^2 + t^2$, and this is at most one when $|x|^2 \le 1 - t^2$. So the ball of radius $\sqrt{1-t^2}$ about zero lies inside E_1. Let s be the coordinate of a point along the line through the marks, measured from their midpoint, and let b be its distance from that line. The square of the product of its distances to the two marks is

$$(b^2 + t^2)^2 + 2s^2(b^2 - t^2) + s^4.$$

If $b^2 > 1 - t^2$, then $b^2 > t^2$ since $t^2 < 1/2$. Then the terms $2s^2(b^2 - t^2)$ and s^4 are nonnegative, so the expression is at least $(b^2 + t^2)^2$, which exceeds one because $b^2 + t^2 > 1$. Thus all of E_1 lies in the cylinder of radius $\sqrt{1-t^2}$ about the line through the marks, and no ball inside that cylinder, wherever its center lies, has a larger radius. This proves $\rho = \sqrt{1-t^2}$.

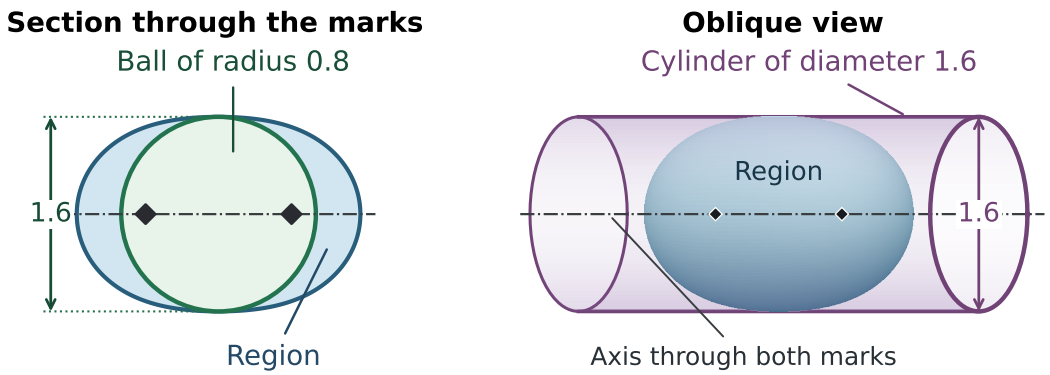


Figure 2: Two marks 1.2 units apart ($t = 0.6$). The region consists of the points whose distances to the two marks have product at most one. A ball of radius 0.8 fits inside it (left). The whole region fits inside a cylinder of diameter 1.6 about the line through the marks (right). Any larger ball would be wider than this cylinder, wherever its center lay, so 0.8 is the largest possible radius.

Jasion and Magnucki [10, (2.2)–(2.4), (3.1)–(3.3)] computed the volume of this region and its meridian curve, that is, its outline in a plane through both marks (Figure 2, left). At threshold one, their dimensionless focal parameter k_c equals our half-separation t. The leading term of the volume loss can also be found from the equation of the boundary in spherical coordinates. For a unit vector θ put $u = \theta \cdot e_1$, and let r be the distance from the origin to the boundary point in the direction θ. Then $(r^2 + t^2)^2 - 4r^2t^2u^2 = 1$. This is a quadratic equation in r^2 with exactly one positive root, namely

$$r^2 = t^2(2u^2 - 1) + \sqrt{1 - 4t^4u^2(1 - u^2)} = 1 + t^2(2u^2 - 1) + O(t^4)$$

uniformly in u. Integrating in spherical coordinates, along each ray from 0 to $r(\theta)$, shows that the volume fraction $|E_1|/\omega_3$ is the mean of $r(\theta)^3$ over the unit sphere. By the expansion above, $r(\theta)^3 = 1 + \frac{3}{2}t^2(2u^2 - 1) + O(t^4)$. The mean of u^2 over the unit sphere is $1/3$, so the mean of $2u^2 - 1$ is $-1/3$, and hence $D = t^2/2 + O(t^4)$. Also $1 - \rho = 1 - \sqrt{1 - t^2} = t^2/2 + O(t^4)$. Thus $(1 - \rho)/D$ tends to one as $t \to 0$, which gives the lower bound for the supremum in Theorem 2.

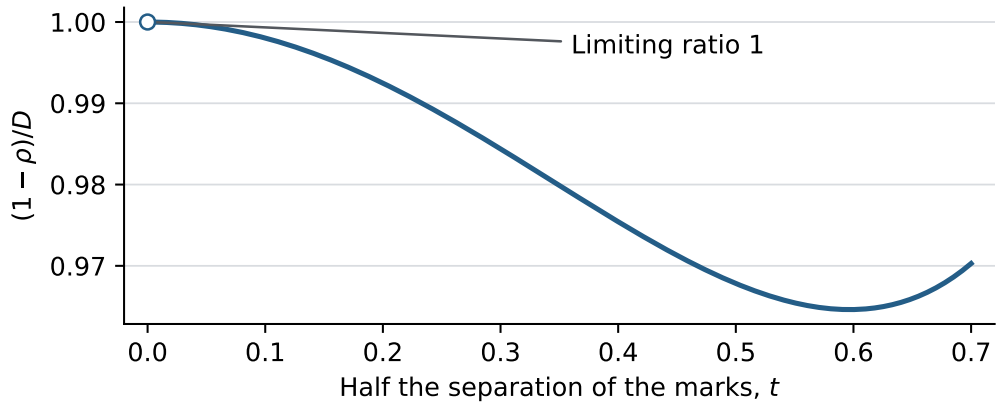


Figure 3: Two equally weighted marks at distance $2t$ from each other, at threshold one. The curve shows the missing fraction of the radius, $1-\rho$, divided by the missing fraction of the volume, D. Both are $t^2/2 + O(t^4)$, so their ratio tends to one as the marks approach each other (open circle).

A center for a large ball inside the region

We now prove the upper bound in Theorem 2. To find a large ball inside E_1 we need a good center. We choose it by minimizing, over all points $z \in \mathbb{R}^3$, the *cost* $\int \log(1+|z-a|^2)\, d\mu(a)$. We write J for the smallest value of the cost:

$$J = \min_{z\in\mathbb{R}^3} \int \log(1+|z-a|^2)\, d\mu(a).$$

The cost is a continuous function of z and tends to infinity as $|z| \to \infty$, so it attains its minimum. Let z be a point where the cost attains its minimum. The gradient of the cost vanishes there, so $\int (z-a)/(1+|z-a|^2)\, d\mu(a) = 0$. Put $M = \int (1+|z-a|^2)^{-1} d\mu(a)$. Jensen's inequality, applied to the exponential function, gives $M \geq e^{-J}$. Now take any v with $|v| \leq 1$, and regard the quotient $|z-a+v|^2/(1+|z-a|^2)$ as a function of the mark a. Since the cost at z equals J, the geometric mean of this quotient with respect to μ is $e^{-J} G_\mu(z+v)^2$. The inequality between the geometric and the arithmetic mean bounds it by the arithmetic mean of the quotient. The numerator equals $|z-a|^2 + 2(z-a)\cdot v + |v|^2$. After division by $1+|z-a|^2$, the middle term integrates to zero, because the gradient vanishes at z. We obtain

$$e^{-J} G_\mu(z+v)^2 \leq \int \frac{|z-a+v|^2}{1+|z-a|^2}\, d\mu(a) = 1 - (1-|v|^2)M.$$

Since $M \geq e^{-J}$, this gives $G_\mu(z+v)^2 \leq e^J - 1 + |v|^2$. The right side is at most one when $|v|^2 \leq 2 - e^J$. So when $J < \log 2$, the closed ball of radius $\sqrt{2-e^J}$ about z is contained in E_1. We now show that nearly maximal volume forces J to be so small that this ball gives $1-\rho \leq D + o(D)$, with the same coefficient one in front of D as in the two-mark example. It remains to prove that, as D tends to zero through positive values,

$$\limsup \frac{J}{\gamma} \leq 2, \qquad \gamma = -\log(1-D). \tag{1}$$

Indeed, (1) forces $J \to 0$, and the ball above then gives $1-\rho \leq 1 - \sqrt{2-e^J} = J/2 + O(J^2)$. Since also $\gamma/D \to 1$, this gives $\limsup (1-\rho)/D \leq 1$, the upper bound in Theorem 2.

The proof of (1) runs along a sequence of distributions, indexed by j, with positive deficits tending to zero. The ball just constructed has its center at z, a point where the cost is smallest. Its radius $\sqrt{2-e^J}$ depends only on the smallest cost J, so we never need to know where z lies. We only need to show that J is small. Since J is the smallest value of the cost, the cost at any point we choose is at least J. If some point has cost $C < \log 2$, then $J \le C$, and the ball about z has radius at least $\sqrt{2-e^C}$. We therefore evaluate the cost at a well-chosen *trial point*. The proof works with a sequence of distributions of marks, indexed by j, and constructs the trial point in two steps. In the first step we change coordinates. We take a compact subset of the region of nearly full volume, and divide all coordinates by c_j, the radius of the ball that has the same Newtonian capacity as this subset. We then choose a point y_j about which the logarithmic moment of the equilibrium measure of the subset is small, and translate y_j to the origin. A point with original position b then has the coordinate $a = b/c_j - y_j$. In these coordinates the equilibrium measures concentrate near the unit sphere. In the second step we choose a point z_j near the new origin about which one quadratic lower bound for the logarithmic moment holds for all marks near the origin. The trial point is the point with coordinate z_j, that is, the point with original position $c_j(y_j + z_j)$.

Equilibrium measure of a nearly maximal region

Let μ_j^0 be any sequence of such distributions with positive deficits $D_j \to 0$, and put $\gamma_j = -\log(1-D_j)$. The jth region has volume $\omega_3(1-D_j) = \omega_3 e^{-\gamma_j}$, which is larger than $\omega_3 e^{-\gamma_j-\gamma_j^2}$. So, by inner regularity of Lebesgue measure, we may choose a compact subset K_j of the jth region with $|K_j| \ge \omega_3 e^{-\gamma_j-\gamma_j^2}$. Let c_j be the radius of the ball with the same Newtonian capacity as K_j, and let ν_j^0 be the Newtonian equilibrium probability measure of K_j. With the kernel $|x-a|^{-1}$, the equilibrium potential is at most $1/c_j$ everywhere [1, Appendix A, Lemma A.1]. Jensen's inequality then gives $\int \log|x-a|\,d\nu_j^0(x) \ge \log c_j$ for every a. The supports of μ_j^0 and ν_j^0 are bounded for each j, and the negative part of $\log r$ is at most $1/r$. The potential bound therefore allows us to exchange the order of integration by Fubini's theorem. Averaging over the marks gives $\log c_j \le \int \log G_{\mu_j^0}\,d\nu_j^0 \le 0$, where the last inequality holds because K_j lies in the region. Thus $c_j \le 1$. On the other hand, the isocapacitary inequality gives $c_j \ge (|K_j|/\omega_3)^{1/3}$. Together with $c_j \le 1$ and the choice of K_j, this gives

$$0 \le \beta_j = -\log c_j \le (\gamma_j + \gamma_j^2)/3.$$

The same argument applies to any compact subset of E_1 of positive volume and shows that its volume is at most ω_3. By inner regularity, $|E_1| \le \omega_3$ for every compactly supported distribution of marks.

Write $\widetilde{\mu}_j, \widetilde{\nu}_j$ for the images of μ_j^0 and ν_j^0 under $x \mapsto x/c_j$. In these coordinates $x^* = x/c_j$ and $a^* = a/c_j$ satisfy

$$|x^*-a^*|^{-1} = c_j|x-a|^{-1}, \qquad \log|x^*-a^*| = \log|x-a| + \beta_j.$$

In the scaled coordinates $\int |x-a|^{-1}\,d\widetilde{\nu}_j(x) \le 1$. Put $h_j(a) = \int \log|x-a|\,d\widetilde{\nu}_j(x)$. Then $h_j \ge 0$ and $\int h_j\,d\widetilde{\mu}_j \le \beta_j$. Since the average of h_j over the marks is at most β_j, there is a point y_j with $h_j(y_j) \le \beta_j + \gamma_j^2$. Let μ_j, ν_j be the images of μ_j^0, ν_j^0 under $T_j(x) = x/c_j - y_j$. Translation gives

$$g_j(a) := \int \log|x-a|\,d\nu_j(x) = h_j(a+y_j), \qquad \int g_j\,d\mu_j = \int h_j\,d\widetilde{\mu}_j \le \beta_j.$$

In particular $g_j(0) = h_j(y_j) \le \beta_j + \gamma_j^2$. The point with coordinate a has original position $T_j^{-1}(a) = c_j(a+y_j)$. The function $\phi(s) = \log s + s^{-1} - 1$ is nonnegative, vanishes only at one, and tends to

infinity at both ends of $(0,\infty)$. Since $\int \phi(|x-a|)\,d\nu_j(x) = g_j(a) + \int |x-a|^{-1}\,d\nu_j(x) - 1$ and the middle integral is at most one, for each a

$$0 \le \int \phi(|x-a|)\,d\nu_j(x) \le g_j(a).$$

At $a=0$ the right side is at most $\beta_j + \gamma_j^2$, which tends to zero. Outside any fixed annulus about the unit sphere, $\phi(|x|)$ is bounded below by a positive constant. The mass of ν_j outside that annulus therefore tends to zero. The sequence (ν_j) is tight. Each weak limit is supported on the sphere, and, because the Newtonian kernel is lower semicontinuous, its potential is at most one everywhere. Its energy is therefore at most one. Among probability measures supported on the unit sphere, the uniform probability σ uniquely minimizes the energy, and its energy is one [1, Appendix A]. Hence ν_j converges weakly to σ.

Logarithmic moments about a common center

The limiting sphere is the unit sphere about the point y_j, which is now the origin. We now construct the point z_j. With z_j we want one lower bound for $g_j(a)$, in terms of the distance from a to z_j, that holds both for marks a near the origin and for distant marks, so that we can average the bound over all the marks. For each sufficiently small fixed $\delta > 0$ we seek points $z_j \to 0$ such that, for every $a \in \mathbb{R}^3$,

$$g_j(a) \ge (1/6 - O(\delta) - o_j(1)) \log(1 + |a - z_j|^2).$$

Here $j \to \infty$ at fixed δ. In the original coordinates, $c_j(y_j + z_j)$ will then be the trial point for J_j, the minimum cost for μ_j^0. The proof has three steps. A strict inequality on the unit sphere gives the bound, uniformly, on each fixed annulus away from zero. A separate estimate treats marks far from the origin. Near the origin a second-derivative estimate selects the point z_j, and the Lipschitz bound for $\log(1+|a|^2)$ lets us move the center of the first two estimates to z_j.

Write $H(s) = \int \log|\theta - se_1|\,d\sigma(\theta)$. For every $s > 0$,

$$H(s) > \tfrac{1}{6}\log(1+s^2). \tag{2}$$

Indeed, regard $H(|x|)$ as a radial function of $x \in \mathbb{R}^3$. Inside the unit sphere its Laplacian is $\int |\theta - se_1|^{-2}\,d\sigma(\theta)$, where $s = |x|$. By Jensen's inequality this is at least $1/(1+s^2)$, since $\int |\theta - se_1|^2\,d\sigma(\theta) = 1+s^2$. The radial Laplacian of the right side of (2) equals $1/(1+s^2) - (2/3)s^2/(1+s^2)^2$. The two functions and their first derivatives vanish at zero, and the Laplacian of H exceeds the other Laplacian by at least $(2/3)s^2/(1+s^2)^2$. Integrating twice, using $\Delta f = s^{-2}(s^2 f')'$, shows that the difference $H(s) - \frac{1}{6}\log(1+s^2)$ is positive and strictly increasing for $0 < s < 1$. The function H is continuous at $s = 1$, by the identity $|\theta - se_1|^2 = (1-s)^2 + s|\theta - e_1|^2$ and the integrability of the logarithm on the sphere. Hence (2) also holds at $s = 1$. For $s > 1$, the identity $H(s) = \log s + H(1/s)$ gives

$$H(s) - \tfrac{1}{6}\log(1+s^2) = H(1/s) - \tfrac{1}{6}\log(1+s^{-2}) + \tfrac{2}{3}\log s > 0.$$

Let (a_j) be a sequence of points, one for each j, that converges to a point $a \neq 0$; these a_j are not the marks of Theorem 1. Then $\liminf_j g_j(a_j) \ge H(|a|)$. Weak convergence alone does not give this, because the logarithm is unbounded below near its pole a_j. So fix $0 < \eta < 1$, write $s = |x - a_j|$, and replace $\log s$ by the truncated function $\max(\log s, \log\eta)$. This function is continuous and bounded below, and the measures ν_j translated by $-a_j$ converge weakly to σ translated by $-a$. Hence the lower limit of the truncated integrals is at least the truncated integral for σ, which is at least $H(|a|)$. The truncation changes the integrand only where $s < \eta$, and there the change is at most $|\log s| \le C/\sqrt{s} \le C\sqrt{\eta}/s$. Since the inverse distance has integral at most one against ν_j, the

truncation raises each integral by at most $C\sqrt{\eta}$, uniformly in j and in the point a_j. Letting η tend to zero proves the claim.

Fix $0 < \delta < A$. For all sufficiently large j we have $g_j(a) \geq \frac{1}{6}\log(1+|a|^2)$ throughout the annulus $\delta \leq |a| \leq A$. Otherwise there would be points a_j of the annulus, for infinitely many j, with $g_j(a_j) < \frac{1}{6}\log(1+|a_j|^2)$. A subsequence would converge to a point a of the annulus, and the lower limit just proved would give $H(|a|) \leq \frac{1}{6}\log(1+|a|^2)$, contrary to (2). The same bound holds for $|a| > A$ when A is large. Indeed $\nu_j(B_2) \to 1$, where B_2 is the radius-two ball about zero. For $|a| > 3$,

$$g_j(a) \geq \nu_j(B_2)\log(|a|-2) - C\sqrt{\nu_j(B_2^c)}.$$

Here the first term comes from B_2, where $|x-a| \geq |a|-2$. The second term bounds the contribution of the complement of B_2: by the Cauchy–Schwarz inequality and $(\log^- s)^2 \leq C/s$, where $\log^- s = \max(0, -\log s)$, the negative part of the logarithm contributes at most $C\sqrt{\nu_j(B_2^c)}$. As $|a| \to \infty$, the ratio $\log(|a|-2)/\log(1+|a|^2)$ tends to $1/2$, which is larger than $1/6$, while $\nu_j(B_2) \to 1$ and the error term tends to zero. This proves the bound for all $|a| \geq \delta$ once j is sufficiently large.

For marks close to zero, we integrate $\phi(|x-a|)$ against ν_j only over the shell $1-\delta \leq |x| \leq 1+\delta$, and we denote this restricted integral by $F_j(a)$. Since $\phi \geq 0$, restricting the integral can only lower it, so $0 \leq F_j \leq g_j$. On this shell the Hessian of the integrand is close to the orthogonal projection onto the radial direction. Averaging these projections over the sphere and applying Taylor's formula will give the coefficient $1/6$. For $|a| \leq \delta < 1/8$ the integrand has Hessian in a equal to $\theta\theta^{\mathsf{T}} + O(\delta)$ in operator norm, with $\theta = x/|x|$. Indeed, put $s = |x-a|$ and $u = (x-a)/s$. The Hessian has eigenvalue $\phi''(s)$ in direction u and $\phi'(s)/s$ in the orthogonal directions. On this shell $|s-1| \leq 2\delta$ and $u = \theta + O(\delta)$. The stated estimate follows from $\phi'(1) = 0$ and $\phi''(1) = 1$. The integral of $\theta\theta^{\mathsf{T}}$ against ν_j over the shell tends to its integral against σ, because ν_j converges weakly to σ and σ gives the boundary of the shell zero mass. Rotational symmetry makes this average a scalar matrix and its trace is one, so the average is $I/3$. The quadratic term of Taylor's formula carries the factor $1/2$, which turns $1/3$ into $1/6$. Consequently, with δ fixed first and j then tending to infinity,

$$\nabla^2 F_j \geq 2b_j I \quad \text{on the ball } |a| \leq \delta, \qquad b_j = \tfrac{1}{6} - O(\delta) - o_j(1) > 0.$$

Choose z_j to minimize F_j on the closed ball $|a| \leq \delta$. Since z_j is a minimizer, the one-sided derivative of F_j at z_j along the segment from z_j to any point a of the ball is nonnegative, even when z_j lies on the boundary of the ball. Taylor's theorem along that segment gives $F_j(a) \geq F_j(z_j) + b_j|a-z_j|^2$, and so, since $g_j \geq F_j$, $g_j(a) \geq F_j(z_j) + b_j|a-z_j|^2$. Since $F_j(z_j) \geq 0$, taking $a = 0$ gives $b_j|z_j|^2 \leq g_j(0) \to 0$. At fixed small δ, b_j stays bounded below by a positive constant, so $z_j \to 0$. Inside the ball, $s^2 \geq \log(1+s^2)$ now gives the desired logarithmic bound about z_j. Outside the ball, put $L(a) = \log(1+|a|^2)$. This function is 1-Lipschitz, so for $|a| \geq \delta$,

$$L(a-z_j) \leq L(a) + |z_j| \leq (1+q_j)L(a), \qquad q_j = \frac{|z_j|}{\log(1+\delta^2)}.$$

Thus, for $|a| \geq \delta$ and all sufficiently large j, the bound $g_j(a) \geq \frac{1}{6}L(a)$ proved above for $|a| \geq \delta$, which we call the *exterior bound*, gives $g_j(a) \geq L(a-z_j)/[6(1+q_j)]$. For fixed δ, $q_j \to 0$. Set

$$k_j = \min\{b_j,\ 1/[6(1+q_j)]\}.$$

The interior and exterior bounds together give, for every a,

$$g_j(a) \geq k_j \log(1+|a-z_j|^2).$$

For fixed small δ the lower limit of k_j as $j \to \infty$ is at least $1/6 - C\delta$, because $b_j = 1/6 - O(\delta) - o_j(1)$ and $q_j \to 0$.

We bound J_j by the cost at the trial point $c_j(y_j + z_j)$. A mark with coordinate a has original position $c_j(a + y_j)$, at distance $c_j|a - z_j|$ from the trial point, and $c_j \le 1$. Therefore

$$J_j \le \int \log(1 + c_j^2 |a - z_j|^2)\, d\mu_j(a) \le k_j^{-1} \int g_j \, d\mu_j \le \beta_j / k_j.$$

Since $\beta_j \le (\gamma_j + \gamma_j^2)/3$, dividing by γ_j and taking the upper limit in j at fixed δ gives $\limsup_j J_j/\gamma_j \le (1/3)/(1/6 - C\delta)$. Letting δ tend to zero gives (1). The ball constructed above shows that the inradius ρ_j for μ_j^0 is at least $\sqrt{2 - e^{J_j}}$ once $J_j < \log 2$. Together with (1) this gives $\limsup_j (1 - \rho_j)/D_j \le 1$. In the example of two approaching marks the ratio $(1 - \rho)/D$ tends to one as $D \to 0$, so for every $h > 0$ the supremum over $0 < D \le h$ is at least one. This proves Theorem 2. □

How the inradius estimate depends on dimension

The stability estimate of Theorem 2 holds in every dimension $d \ge 3$, with the coefficient $1/(d-2)$ in place of one. The example of two marks already shows why the coefficient changes: the inradius of the region of two marks does not depend on the dimension, but the fraction of the volume that is missing does.

Place equal masses at $\pm t e_1$ with $0 < t < 1/\sqrt{2}$ and take $R = 1$. Every plane that contains the line through the marks cuts the region in the same planar shape, whatever the dimension. As in Figure 2, a ball of radius $\sqrt{1 - t^2}$ fits inside the region, and the region fits inside a cylinder of that radius about the line through the marks. So the inradius is $\sqrt{1 - t^2}$ in every dimension $d \ge 2$. On each ray from the midpoint of the marks, the equation of the boundary is a quadratic equation in r^2 with leading coefficient one and constant term $t^4 - 1 < 0$. It has exactly one positive root $r(\theta)^2$. Thus on each ray the region occupies $0 \le r \le r(\theta)$, and the volume fraction is the spherical mean of $r(\theta)^d$.

Theorem 3. *Fix $d \ge 3$ and $R > 0$. For a compactly supported probability measure μ on $\mathbb{R}^d$, define G_μ, E_R, ρ as in Theorem 2, and put $D = 1 - |E_R|/(\omega_d R^d)$. Then*

$$\lim_{h \downarrow 0} \sup_{\mu:\, 0 < D \le h} \frac{1 - \rho/R}{D} = \frac{1}{d-2}.$$

The same limit holds if the supremum is taken only over finite sets of equally weighted marks.

Proof. By rescaling we may take $R = 1$. Write $k = d - 2$. The argument that produced the inscribed ball before (1) uses only the Euclidean inner product and Jensen's inequality, so it holds in $\mathbb{R}^d$ as well: with $J = \min_z \int \log(1 + |a - z|^2)\, d\mu(a)$, it again gives $\rho \ge \sqrt{2 - e^J}$ whenever $J < \log 2$. Since $1 - \sqrt{2 - e^J} = J/2 + O(J^2)$, it suffices to prove that $\limsup J/\gamma \le 2/k$, where $\gamma = -\log(1 - D)$. This gives the coefficient $1/k$ for the loss of radius.

Take any sequence μ_j^0 of compactly supported distributions of marks with positive deficits $D_j \to 0$. Write $\gamma_j = -\log(1 - D_j)$, and let J_j be the smallest cost for μ_j^0. Choose a compact set $K_j \subset E_1$ of volume at least $\omega_d e^{-\gamma_j - \gamma_j^2}$. Let c_j be the radius of the ball with the same Newtonian capacity as K_j, and let ν_j^0 be the Newtonian equilibrium probability measure of K_j. With the kernel $|x - a|^{-k}$, the potential of ν_j^0 is at most c_j^{-k} everywhere [1, Appendix A]. By Jensen's inequality,

$$-k \int \log|x - a|\, d\nu_j^0(x) \le \log \int |x - a|^{-k}\, d\nu_j^0(x) \le -k \log c_j.$$

Hence $\int \log|x-a|\,d\nu_j^0(x) \ge \log c_j$ for every a. Averaging this inequality over the marks gives $\log c_j \le \int \log G_{\mu_j^0}\,d\nu_j^0 \le 0$, where the last inequality holds because K_j lies in E_1. As in dimension three, the bound on the potential controls the negative part of the logarithm and justifies the exchange of the two integrals. By the isocapacitary inequality,

$$0 \le \beta_j = -\log c_j \le (\gamma_j + \gamma_j^2)/d.$$

We use the same coordinates as in the three-dimensional proof. Let $\widetilde{\mu}_j, \widetilde{\nu}_j$ be the images of μ_j^0 and ν_j^0 under the dilation $x \mapsto x/c_j$. At every point a the potential of $\widetilde{\nu}_j$ with kernel $|x-a|^{-k}$ is at most one, and the logarithmic moment $h_j(a) = \int \log|x-a|\,d\widetilde{\nu}_j(x)$ is nonnegative. The average of h_j over the marks satisfies $\int h_j\,d\widetilde{\mu}_j \le \beta_j$, so there is a point y_j with $h_j(y_j) \le \beta_j + \gamma_j^2$. Define μ_j, ν_j as the images of μ_j^0, ν_j^0 under $T_j(x) = x/c_j - y_j$. In these coordinates

$$g_j(a) := \int \log|x-a|\,d\nu_j(x) = h_j(a+y_j), \qquad \int g_j\,d\mu_j = \int h_j\,d\widetilde{\mu}_j \le \beta_j.$$

Thus $g_j(0) \le \beta_j + \gamma_j^2$. The point z_j chosen below gives the trial point $T_j^{-1}(z_j) = c_j(z_j + y_j)$ for the original distribution. As in dimension three, the cost at the trial point is only an upper bound for the smallest cost J; the inscribed ball itself is centered at a point where the cost is smallest.

The function

$$\phi_k(s) = \log s + \frac{s^{-k}-1}{k}$$

is nonnegative, vanishes only at $s=1$, and tends to infinity at both ends of $(0,\infty)$. For every a we have $\int \phi_k(|x-a|)\,d\nu_j(x) \le g_j(a)$. At $a=0$ the right side is at most $\beta_j + \gamma_j^2$, which tends to zero. Outside any fixed annulus about the unit sphere, $\phi_k(|x|)$ is bounded below by a positive constant, so the mass of ν_j outside that annulus tends to zero. Every weak limit of ν_j is therefore supported on the unit sphere. Because the kernel $|x-a|^{-k}$ is lower semicontinuous, the potential of the limit is at most one everywhere, and hence its energy is at most one. Among probability measures supported on the unit sphere, the uniform probability σ uniquely minimizes this energy, with value one [1, Appendix A]. Therefore ν_j converges weakly to σ.

Put $\alpha = k/(2d)$. We claim that, for any sufficiently small fixed $\delta > 0$, there is a point z_j such that for every a

$$g_j(a) \ge (\alpha - O_d(\delta) - o_j(1)) \log(1 + |a - z_j|^2). \tag{3}$$

Here $o_j(1) \to 0$ with δ held fixed. For $|a| \ge \delta$ we need the bound for all such a at once, however far away they are. For $|a| \le \delta$ a second-derivative estimate will determine the center z_j.

First let $H(s) = \int \log|\theta - se_1|\,d\sigma(\theta)$, with e_1 the first coordinate unit vector. For $0 < s < 1$ the radial Laplacian of H is $k\int |\theta - se_1|^{-2}\,d\sigma(\theta) \ge k/(1+s^2)$. The radial Laplacian of $\alpha\log(1+s^2)$ equals

$$\frac{k}{1+s^2} - \frac{2k}{d}\frac{s^2}{(1+s^2)^2}.$$

Both functions and their first radial derivatives vanish at zero. Integrating the difference of the radial Laplacians twice shows that $H(s) - \alpha\log(1+s^2)$ is positive and strictly increasing on $(0,1)$. The function H is continuous at $s=1$, because the logarithmic singularity is integrable on the sphere; this follows from $|\theta - se_1|^2 = (1-s)^2 + s|\theta - e_1|^2$. For $s > 1$ we use the identity $H(s) = \log s + H(1/s)$: the difference $H(s) - \alpha\log(1+s^2)$ equals the same difference at $1/s$ plus $(1-2\alpha)\log s$, and both terms are positive. Thus $H(s) > \alpha\log(1+s^2)$ for every $s > 0$.

If $a_j \to a \ne 0$, then $\liminf g_j(a_j) \ge H(|a|)$. As in dimension three, this follows by truncating the logarithm and taking weak limits. Indeed for $0 < s < \eta < 1$ we have $|\log s| \le C_k s^{-k/2} \le C_k \eta^{k/2} s^{-k}$,

so the truncation changes each integral by at most $C_k \eta^{k/2}$, uniformly in j and in the center. Since $H(s) > \alpha \log(1+s^2)$ and the annulus is compact, the argument by contradiction used in dimension three gives $g_j(a) \geq \alpha \log(1+|a|^2)$ on any fixed annulus $\delta \leq |a| \leq A$ for large j. Now let B_2 be the ball of radius two about zero. If $|a| > A > 3$, then $|x-a| \geq |a| - 2$ for x in B_2, and splitting the integral over B_2 and its complement gives

$$g_j(a) \geq \nu_j(B_2) \log(|a|-2) - C_k \sqrt{\nu_j(B_2^c)}.$$

The last term comes from the Cauchy–Schwarz inequality and $(\log^- s)^2 \leq C_k s^{-k}$, as in dimension three. Since $\nu_j(B_2) \to 1$ and $\log(s-2)/\log(1+s^2) \to 1/2 > \alpha$, one can first choose A and then j large enough that $g_j(a) \geq \alpha \log(1+|a|^2)$ for every $|a| > A$.

Next restrict $\int \phi_k(|x-a|)\, d\nu_j(x)$ to $1-\delta \leq |x| \leq 1+\delta$ and call the result $F_j(a)$. For $|a| \leq \delta < 1/8$, its integrand has Hessian $k\theta\theta^{\mathsf{T}} + O_d(\delta)$, where $\theta = x/|x|$. Indeed the Hessian at distance $s = |x-a|$ has radial eigenvalue $\phi_k''(s)$ and tangential eigenvalues $\phi_k'(s)/s$. The projection onto the radial line differs from $\theta\theta^{\mathsf{T}}$ by $O(\delta)$ in operator norm, and $\phi_k'(1) = 0$, $\phi_k''(1) = k$. The annular integral of $\theta\theta^{\mathsf{T}}$ tends to I/d by rotational symmetry of σ and weak convergence. Thus $\nabla^2 F_j \geq 2b_j I$ with $b_j = \alpha - O_d(\delta) - o_j(1) > 0$ for fixed small δ and large j. Let z_j be a point where F_j is smallest on the closed ball $|a| \leq \delta$. Since z_j is a minimizer, the one-sided derivative of F_j at z_j along the segment to any point a of the ball is nonnegative, and Taylor's theorem along that segment gives $g_j(a) \geq F_j(z_j) + b_j |a - z_j|^2$ throughout the ball. Taking $a = 0$ gives $b_j |z_j|^2 \leq g_j(0) \to 0$, so $|z_j| \to 0$. Since $s^2 \geq \log(1+s^2)$, this proves (3) for $|a| \leq \delta$. For $|a| \geq \delta$, the function $L(a) = \log(1+|a|^2)$ is 1-Lipschitz and $L(a) \geq \log(1+\delta^2)$, so $L(a - z_j) \leq L(a) + |z_j| \leq [1 + |z_j|/\log(1+\delta^2)]\, L(a)$. Together with the bound $g_j(a) \geq \alpha L(a)$ for $|a| \geq \delta$, this proves (3) there as well.

Write b for the positive coefficient on the right of (3) at fixed j, δ. Returning to the original coordinates multiplies all distances by $c_j \leq 1$. The cost at the trial point $c_j(y_j + z_j)$ therefore gives

$$J_j \leq \int \log(1 + c_j^2 |a - z_j|^2)\, d\mu_j(a) \leq b^{-1} \int g_j \, d\mu_j \leq \beta_j / b.$$

Divide by γ_j, take $j \to \infty$, and then let δ tend to zero. This gives $\limsup J_j/\gamma_j \leq 1/(d\alpha) = 2/k$, and the bound $\rho_j \geq \sqrt{2 - e^{J_j}}$ then gives $\limsup (1-\rho_j)/D_j \leq 1/k$.

For the reverse bound, put equal masses at $\pm t e_1$. As shown before the theorem, $\rho = \sqrt{1-t^2}$ for $0 < t < 1/\sqrt{2}$ in every dimension $d \geq 2$. For a unit vector θ put $u = \theta \cdot e_1$. The equation of the boundary is the same as in dimension three and gives $r^2 = 1 + (2u^2 - 1)t^2 + O(t^4)$ uniformly in u. The volume fraction is now the spherical mean of $r^d = (r^2)^{d/2} = 1 + \frac{d}{2}(2u^2 - 1)t^2 + O_d(t^4)$, and the mean of u^2 is $1/d$. Thus $D = (d-2)t^2/2 + O_d(t^4)$, whereas $1 - \rho = t^2/2 + O(t^4)$. The ratio $(1-\rho)/D$ therefore tends to $1/(d-2)$ as $t \to 0$. So for every $h > 0$ the supremum is at least $1/(d-2)$, even when it is taken only over finite sets of equally weighted marks. □

At threshold one, Theorem 3 also gives the sharp lower bound for the inradius when the loss of volume is prescribed. Write $I_d(\varepsilon)$ for the infimum of ρ over all distributions of marks with $0 \leq D \leq \varepsilon$. We show that $I_d(\varepsilon) = 1 - \varepsilon/(d-2) + o(\varepsilon)$ as ε tends to zero.

First let $D = 0$. Keep the distribution fixed, and choose compact subsets $K_j \subset E_1$ whose volumes tend to ω_d from below. In the proof above we may take $1/j$ as the tolerance in the choice of K_j, so that $0 \leq \beta_j \leq (1/j + 1/j^2)/d$. The estimate $J \leq \beta_j / b$, in which b stays bounded away from zero for each fixed small δ, then gives $J = 0$. The integrand in the definition of J is zero only at the minimizing center and positive everywhere else. So the whole distribution sits at that point: it is a point mass, and $\rho = 1$.

Now let $0 < D \leq \varepsilon$. By Theorem 3, the loss $1 - \rho$ is at most $(1/(d-2) + o(1))\varepsilon$. Two equally weighted marks attain this loss to first order, for every sufficiently small $\varepsilon > 0$. Indeed, by the

radial volume formula their deficit $D(t)$ is a continuous function of t. It is zero at $t = 0$ and exceeds ε at $t = 2\sqrt{\varepsilon/(d-2)}$, so $D(t) = \varepsilon$ for some t in between. The two expansions above then give $1-\rho = \varepsilon/(d-2)+O(\varepsilon^2)$, with a constant that depends only on d. Hence $I_d(\varepsilon) = 1-\varepsilon/(d-2)+o(\varepsilon)$.

Two marks in the plane show why Theorem 3 requires $d \geq 3$. In the plane, write a point in polar coordinates about the midpoint of the marks, with $u = \cos\theta$. The boundary of the region of two marks is then given by $r^2 = t^2\cos(2\theta) + \sqrt{1 - t^4\sin^2(2\theta)}$. The mean of r^2 over the angle θ is the area fraction, hence $D = t^4/4 + O(t^8)$, while $1 - \rho = t^2/2 + O(t^4)$. Thus $1 - \rho \sim \sqrt{D}$ as t tends to zero. So already for two marks in the plane, the fractional loss of inradius is not bounded by a constant times the fractional loss of area.

Which distance means satisfy the volume bound?

The geometric mean in Erdős's question belongs to a one-parameter family. Given the finite marks and positive weights of Theorem 1, define

$$M_p(x) = \begin{cases} \left(\sum_j w_j|x - a_j|^p\right)^{1/p}, & p \neq 0, \\ \prod_j |x - a_j|^{w_j}, & p = 0. \end{cases}$$

For $p < 0$ we set $M_p = 0$ at a mark, which is the limit of $M_p(x)$ as x approaches the mark. For fixed x, the weighted power mean $M_p(x)$ is nondecreasing in p. We show that $2 - d$, the exponent of the Newtonian kernel, is the smallest exponent for which the volume bound of Theorem 1 holds for every configuration. In three dimensions this exponent is -1. The geometric mean ($p = 0$), which is the mean in Erdős's question, satisfies the volume bound. For $p < -1$, two equally weighted marks that are sufficiently close give, at $R = 1$, a volume greater than that of the unit ball.

Theorem 4. *Fix $d \geq 3$ and $p \in \mathbb{R}$. The inequality*

$$|\{M_p \leq R\}| \leq \omega_d R^d$$

holds for all finite configurations, positive weights of sum one and $R > 0$ if and only if $p \geq 2 - d$. In this range, for each $R > 0$, equality holds exactly when all the marks coincide. For every $p < 2 - d$, two equally weighted marks suffice to violate the inequality.

Consequences of earlier inequalities

For $-d < p < 0$ put $s = d + p$; thus p is the exponent of the distance and s is the order of the Riesz potential. For $2 \leq s < d$, the sharp weak estimate of Tang and Wang [8, Theorem 1.1] gives $|\{V > \lambda\}| \leq \omega_d \lambda^{-d/(d-s)}$ for $\lambda > 0$, where V is the potential of a probability density with the kernel $|x|^{s-d}$, without a normalizing constant. To apply the estimate to finitely many marks, replace the point mass at each mark a_j by w_j times a smooth probability density supported in a small ball about a_j, and let the radii of the balls tend to zero. The sum of these densities is a probability density. Away from the marks, the resulting potentials converge to the potential of the marks, $V_{\mathrm{at}}(x) = \sum_j w_j|x - a_j|^p$. By Fatou's lemma, the same bound holds for the strict superlevel sets $\{V_{\mathrm{at}} > \lambda\}$. As positive thresholds increase to λ, these strict superlevel sets decrease to $\{V_{\mathrm{at}} \geq \lambda\}$, and their volumes are finite, so the bound holds for $\{V_{\mathrm{at}} \geq \lambda\}$ as well. For $s = 2$ this is the volume bound for M_{2-d}. Since M_p is nondecreasing in p, the bound follows for every $p \geq 2 - d$.

For $2 - d < p < 0$, the statement that equality forces the marks to coincide also follows from Clark and Laugesen [1, Theorem 1.1(iii)–(iv)]. Let $K = \{M_p \leq R\}$, let c be the radius of the

ball with the same Newtonian capacity as K, and let ν be the Newtonian equilibrium probability measure of K. On K we have $\sum_j w_j|x-a_j|^p \geq R^p$, because $p<0$. Part (iii), applied at each mark, gives $\int_K |x-a_j|^p\,d\nu(x) \leq c^p$. Hence

$$R^p \leq \sum_j w_j \int_K |x-a_j|^p\,d\nu(x) \leq c^p.$$

Thus $c \leq R$. If the volume of K equals $\omega_d R^d$, then $c=R$ by the isocapacitary inequality, and each of the inequalities from part (iii) is an equality. By part (iv), K then contains the whole sphere of radius R about each mark. Choose a unit vector u, and among the marks choose one, a, for which $u\cdot a$ is largest. At the point $a+Ru$, every distance to a mark is at least R, and the distance to any mark at a different point is greater than R. Since $p<0$, if the marks do not all coincide then $\sum_j w_j|a+Ru-a_j|^p < R^p$, so $a+Ru$ is not in K. This contradicts the fact that K contains the sphere of radius R about a. Finally, for $p' \geq p$ we have $M_{p'} \geq M_p$, so $\{M_{p'} \leq R\}$ is contained in $\{M_p \leq R\}$, and the strict volume bound for marks that do not all coincide holds for all larger exponents as well.

At $p=2-d$, write $\lambda = R^{2-d}$ and $U(x)=\sum_j w_j|x-a_j|^{2-d}$, with $U=+\infty$ at each mark. Let $K=\{U\geq\lambda\}$ and $\Omega=\{U>\lambda\}$. The inequality of part (iii) includes $p=2-d$, so the ball with the same capacity as K has radius at most R. At this exponent, however, equality in part (iii) at a mark says only that the equilibrium potential of K at the mark equals the equilibrium energy of K. Each mark is an interior point of K, and the equilibrium potential takes this value at every interior point of K. So this equality holds at every mark even when the marks are distinct, and it cannot by itself force the marks to coincide. To show that the marks coincide we use instead the equality case of the isocapacitary inequality [2, §1.1]. The level set $\{U=\lambda\}$ has volume zero, because U is analytic away from the marks. If $|K|=\omega_d R^d$, then $|\Omega|=\omega_d R^d$ as well. Let B_R be a ball of radius R. The isocapacitary inequality and the monotonicity of Newtonian capacity give

$$\operatorname{cap}(B_R) \leq \operatorname{cap}(\Omega) \leq \operatorname{cap}(K) \leq \operatorname{cap}(B_R).$$

Hence, by the equality case, Ω agrees with a ball of radius R up to a set of volume zero. This argument applies even when λ is a critical value of U.

To see why this determines the marks, suppose more generally that $\{U>\lambda\}$ agrees with an open ball $B(z,b)$ up to a set of volume zero. Then every mark lies in the open ball. Indeed, a neighborhood of each mark lies in $\{U>\lambda\}$, whereas every neighborhood of a point on or outside the sphere $|x-z|=b$ contains a set of positive volume outside the ball. Next, $U=\lambda$ at each point of the sphere. Indeed, if $U>\lambda$ at such a point, then by continuity $U>\lambda$ on a neighborhood of it, which contains a set of positive volume outside the ball; and if $U<\lambda$ there, then $U<\lambda$ on a neighborhood, which contains a set of positive volume inside the ball. Since all the marks are inside the ball, U is harmonic outside the ball, equals λ on the sphere, and tends to zero at infinity. By uniqueness for the exterior Dirichlet problem,

$$U(x) = \lambda\, b^{d-2}|x-z|^{2-d} \qquad (|x-z|>b). \tag{4}$$

Both sides are harmonic on the complement of z and the marks, which is connected, so by analytic continuation the identity holds on that whole set. If a mark lay at a point other than z, the left side would tend to infinity at that mark, because its weight is positive, while the right side would stay bounded. Thus all marks coincide at z. Taking $b=R$ completes the deduction.

Now let $p<2-d$ and put $q=-p>d-2$. Consider n equally weighted marks $a_1,\dots,a_n$ in the closed unit ball. For each placement of the marks, take the minimum over all points x of that ball

of $n^{-1}\sum_j |x-a_j|^{-q}$, and let P_n be the supremum of this minimum over all placements. Erdélyi and Saff [9, Theorem 3.2] show that $P_n \to \infty$ when $q \geq d$, and that for $d-2 < q < d$ the limit of P_n is

$$\frac{\Gamma((d-q)/2)\Gamma(1+q/2)}{\Gamma(d/2)} > 1.$$

To see that this quotient exceeds one, note that it equals one at $q = d-2$ and that its derivative in q is positive on $d-2 < q < d$, because the logarithmic derivative of Γ is strictly increasing and $1+q/2 > (d-q)/2$. Thus in both ranges there are a number n and a placement of n equally weighted marks for which the minimum above is some number $P > 1$. For this placement, $M_{-q} \leq P^{-1/q}$ throughout the unit ball. Put $R = P^{-1/q} < 1$. Then the set $\{M_{-q} \leq R\}$ contains the unit ball, so its volume is at least ω_d, which is greater than $\omega_d R^d$.

Direct flux proof and two-mark counterexamples. Use the Newtonian potential U, the superlevel volume $W(\lambda)$, and the equal-volume radius $r(\lambda)$ from the flux proof of Theorem 1. That argument proves that $r(\lambda)^{2-d} - \lambda$ is nondecreasing and has limit zero as λ tends to zero. At the endpoint $p = 2-d$ we have $M_{2-d} = U^{1/(2-d)}$, so $\{M_{2-d} \leq R\} = \{U \geq R^{2-d}\}$. To prove the strict inequality at this exponent we examine when equality holds in the isoperimetric inequality.

Suppose that $\lambda > 0$ is a regular value of U and that equality holds in the isoperimetric inequality for $\{U > \lambda\}$. Then $\{U > \lambda\}$ is a ball up to a set of volume zero. By the identity (4) and the argument that follows it, every mark then lies at the center of that ball.

Consequently, if the marks do not all coincide, the isoperimetric inequality is strict at every regular value. It follows that $(r^{2-d})' > 1$ almost everywhere on $(0,\infty)$. By local absolute continuity, $r(\lambda)^{2-d} - \lambda$ is strictly increasing there. Its limit at zero is zero, so $r(\lambda)^{2-d} > \lambda$ for every $\lambda > 0$. Set $\lambda = R^{2-d}$ to obtain $|\{M_{2-d} \leq R\}| < \omega_d R^d$. Since $\{M_p \leq R\}$ is contained in $\{M_{2-d} \leq R\}$ for $p \geq 2-d$, the same strict bound holds for every $p \geq 2-d$. If all marks coincide at a, then $M_p(x) = |x-a|$ and the set $\{M_p \leq R\}$ is the ball of radius R about a.

For the converse, fix $p < 2-d$, take $R = 1$, and put equal weights at $\pm te_1$ with $0 < t < 1/2$. By the triangle inequality, the set $\{M_p \leq 1\}$ contains the ball of radius $1-t$ and lies in the ball of radius $1+t$, both about the origin. On each ray from the origin, both distances to the marks increase strictly once the distance from the origin exceeds t. Each ray therefore meets the boundary of the set at exactly one point, at a distance $r = r(\theta)$ from the origin with $1-t \leq r \leq 1+t$. For the direction θ put $u = \theta \cdot e_1$, and put $h = t/r$. The mean of the pth powers of the two distances at the point $r\theta$ is

$$\frac{r^p}{2}\bigl((1-2uh+h^2)^{p/2} + (1+2uh+h^2)^{p/2}\bigr) = r^p\left[1 + \frac{p}{2}\{1+(p-2)u^2\}h^2 + O_p(h^4)\right].$$

The terms with odd powers of h cancel, because the two brackets differ only in the sign of uh; the error is uniform in $|u| \leq 1$ for small t. At $t = 0$ the derivative of the left side with respect to r at $r = 1$ equals $p \neq 0$. So we can solve the boundary equation, in which the left side equals one, and obtain, uniformly in the direction,

$$r = 1 - \tfrac{1}{2}\{1+(p-2)u^2\}t^2 + O_p(t^4).$$

Averaging r^d over the sphere, and using that the spherical mean of u^2 is $1/d$, gives

$$\frac{|\{M_p \leq 1\}|}{\omega_d} = 1 - \tfrac{1}{2}(d+p-2)t^2 + O_{p,d}(t^4) > 1$$

for all sufficiently small positive t. On the annulus $1-t \leq r \leq 1+t$ both distances to the marks are at least $1-2t > 0$, so the expansion is valid for each fixed $p < 2-d$. □

Discussion

Extending Pólya's planar theorem, we proved that in every dimension $d \geq 3$ the largest possible volume of the region is that of the unit ball, attained only when all marks coincide. We generalized this result to weighted marks. Our proof corrects Erdős's 1976 report of a two-point counterexample in three dimensions by showing that every pair of distinct marks gives strictly less volume than the unit ball. In every dimension $d \geq 3$ we also showed that, as the volume approaches its maximum, the worst possible fractional loss of inradius agrees to first order with $1/(d-2)$ times the fractional loss of volume; in three dimensions this coefficient is one. Finally, we replaced the geometric mean by the weighted power mean of order p of the distances, and showed that the volume bound then holds for all marks and weights exactly when $p \geq 2 - d$.

These results suggest several further questions. First, fix a dimension $d \geq 3$ and let ε tend to zero. Among regions at threshold one with at least $(1 - \varepsilon)$ times the volume of the unit ball, the sharp lower bound for the inradius is $1 - \varepsilon/(d-2) + o(\varepsilon)$. This is Theorem 2 when $d = 3$ and Theorem 3 in general. What is the next term of this expansion? Second, in each such dimension, two equally weighted marks approaching one another attain this first-order rate. Which other arrangements of marks attain it? Third, our estimate fails in the plane. Already for two marks there, the fractional loss of inradius is asymptotically the square root of the fractional loss of area, which is much larger, as we saw after the proof of Theorem 3. For a fixed number of equally weighted marks in the plane, what is the sharp bound on the fractional loss of inradius as the fractional loss of area tends to zero?

Acknowledgments

This work was done overwhelmingly by AI models, steered by the author in a continuous critical dialogue, alternating between Anthropic's Claude and OpenAI's ChatGPT. The author takes full responsibility for the paper.

References

[1] C. Clark and R. S. Laugesen, *Balls minimize moments of logarithmic and Newtonian equilibrium measures*, Potential Anal. **63** (2025), 389–412, published online in 2024.

[2] G. De Philippis, M. Marini and E. Mukoseeva, *The sharp quantitative isocapacitary inequality*, Rev. Mat. Iberoam. **37** (2021), 2191–2228.

[3] P. Erdős, *Extremal problems on polynomials*, in *Approximation Theory II*, Academic Press (1976), 347–355.

[4] D. Goel, *Erdős's 1974 distance-product volume problem resolved: the ball is always the unique maximizer in $d \geq 2$*, preprint (2026).

https://doi.org/10.5281/zenodo.24378921

[5] W. K. Hayman and E. F. Lingham, *Research Problems in Function Theory (New Edition)*, arXiv:1809.07200 (2018), Problem 4.23 and Tables 1–2.

[6] A. Ferudun, *On a question of Erdős on lemniscates in $\mathbb{R}^m$: connected lemniscates lie in a ball of radius 2*, preprint (2026).

[7] G. Talenti, *Elliptic equations and rearrangements*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **3** (1976), 697–718.

[8] H. Tang and Y. Wang, *Sharp constants for weak estimates of the Riesz potentials*, arXiv:2608.31043v2 (2026).

[9] T. Erdélyi and E. B. Saff, *Riesz polarization inequalities in higher dimensions*, J. Approx. Theory **171** (2013), 128–147.

[10] P. Jasion and K. Magnucki, *Elastic buckling of Cassini ovaloidal shells under external pressure – theoretical study*, Arch. Mech. **67** (2015), 179–192.

[11] P. Erdős, F. Herzog and G. Piranian, *Metric properties of polynomials*, J. Analyse Math. **6** (1958), 125–148.

[12] G. Pólya, *Beitrag zur Verallgemeinerung des Verzerrungssatzes auf mehrfach zusammenhängende Gebiete (Zweite Mitteilung)*, Sitzungsber. Preuss. Akad. Wiss., Phys.-Math. Kl. (1928), 280–282.